Submission in Response to NSF CI 2030 Request for Information

DATE AND TIME: 2017-04-04 16:36:25
REFERENCE NO: 217

Author Names & Affiliations

- Gregory Monaco - Kansas State University & The Great Plains Network
- Donald (Rick) McMullen - Texas A&M University

Contact Email Address (for NSF use only)

(Hidden)

Research Domain, discipline, and sub-discipline

Experimental Psychology, Chemistry, Computer Networking, High Performance Computing, Cyberinfrastructure

Title of Submission

Organizing to broaden access to national, regional and campus-based cyberinfrastructure

Abstract (maximum ~200 words).

While the development and deployment of new cyberinfrastructure technologies and facilities to meet emerging research needs is of primary importance, a no less important need is to organize for effective uptake and use of these CI tools as broadly as possible. We recommend a program that leverages existing organizations to create a feedback loop between users and providers of CI that is responsive both broadly across disciplines and deeply to faculty and students at less well-resourced campuses who represent the next generation of innovation in the United States. This idea is expanded below and potential solutions are outlined.

Question 1 Research Challenge(s) (maximum ~1200 words): Describe current or emerging science or engineering research challenge(s), providing context in terms of recent research activities and standing questions in the field.

Development and deployment of new CI technologies and facilities to meet emerging research needs are clearly on the forefront of this call for papers. A second important need is to organize for effective uptake and use of CI as broadly as possible. There is an evolving ecosystem of support services to meet the need for uptake and use of CI that leverages existing organizations to create a feedback loop between users and providers of CI that is responsive to the broad science and engineering community.

Innovation and discovery are influenced by organization and the availability of accessible technologies. Creating the conditions for innovation and discovery is a research challenge in itself, particularly with regard to the diffusion of both access to technologies and expertise in their use. A primary challenge is building researchers' understanding of computing and data literacy as tools for science. During preparation of students as researchers, access to CI is a critical part of their education and training environment. A second challenge is the need to provide cyberinfrastructure, including expertise, to faculty and students at at smaller schools and minority serving institutions with

less access to national research CI resources.

Several organizational studies have been conducted of national, centralized CI projects such as the TeraGrid and XSEDE [1,2], but no studies have been done on the complex set of CI providers that spans national resources to researchers residing on academic campuses. The research need is to understand how this ecosystem works and how it can be purpose-built, or at least directed and tuned to meet the needs of a broader set of researchers and students. In particular, based on the outcomes of the “Role of Regional Organizations” conference [3], there is a definite role for organizations with regional scope such as regional optical networks and statewide research computing consortia, or federations of these, to provide the following:

- Broadening the uptake and use of national CI resources by smaller colleges and universities, MSIs, tribal colleges and research facilities with a regional focus (e.g. regional water, agriculture, and natural resources research centers)

- Organizing the sharing and reuse of existing and future NSF investments in CI on campuses (campus-campus bridging)

- Coordinating the training and professional development of CI experts, with a focus on meeting regional/local knowledge and expertise needs, and supporting research in a regional context

- Coordination of cooperative development of CI across many institutions within a region

References
1. N. Berente, J. Howison, J. L. King, K. J. Lyytinen, and N. Wilkins-Diehr, “Virtual Organizations as Sociotechnical Systems: Exploring How Organization Scientists and Virtual Organization Leaders Can Collaborate,” 2013.

2. J. Cummings, T. Finholt, I. Foster, C. Kesselman, and K. A. Lawrence, “Beyond being there: A blueprint for advancing the design, development, and evaluation of virtual organizations,” 2008.

3. Monaco, G.E., McMullen, D.F., Huntoon, G., Leasure, J., Swanson, D., Neeman, H., Blake, J., Adams, K. The Role of Regional Organizations in Improving Access to the National Computational Infrastructure: A Report to the National Science Foundation, June, 2016.

Question 2 Cyberinfrastructure Needed to Address the Research Challenge(s) (maximum ~1200 words): Describe any limitations or absence of existing cyberinfrastructure, and/or specific technical advancements in cyberinfrastructure (e.g. advanced computing, data infrastructure, software infrastructure, applications, networking, cybersecurity), that must be addressed to accomplish the identified research challenge(s).

The report from the Role of Regional Organizations [1] outlined the needs as follows:

“If the challenge is to expand discovery and innovation via newly developed CI tools, then a cadre of organizations and people who are close to the end user community of scientists must be empowered and understand their role in terms of expanding access to these tools to the broader community.”

The report made the following four recommendations (among others) for developing CI at a regional level and for building regional CI organizations into an end-to-end ecosystem view of research computing:

1. “Support acquisition of regional cyberinfrastructure resources to serve a geographic region where the performing regional organization shows commitment to some or all of the following: Operation of the CI resource (including funding the system administration staff, operations cost of the resource and network connectivity); user support, outreach, user training and education; integration of the system with national CI (e.g. level two XSEDE provider and OSG[/Jetstream] partner, and other regional CI); low barrier of entry for regional researchers to increase and broaden participation; integration of the instrument with high-performance networking and advanced data transfer and data sharing capabilities.
Note: Many university consortia that might compete to acquire and operate regional CI resources on behalf of multiple members are dissuaded from doing so when they are administered by a member university. The consortium must, first, compete on a member campus to acquire a limited submission slot in response to infrastructure solicitations (e.g., NSF’s Major Research Instrumentation competition). At

one time, consortia were exempt from limitations and did not compete for slots on their own campus. This made it easier for funding agencies to award funds for resources to regional consortia (with bylaws & clearly executed membership agreements) on behalf of their entire membership."

2. "Fund regional centers of CI excellence that are committed to the principles of sharing, education and outreach. Regional centers of CI excellence would assist with sharing resources at multiple levels, with engaging underserved groups and with increasing educational opportunities in the region. These centers would also provide a liaison function between campuses and national infrastructure. Successful models are starting to emerge. However, rather than cloning existing models it is essential that regional CI centers are unique to and meet the needs of a geographic area."

3. "Develop a policy to require a resource-sharing plan akin to a data management plan."

4. "Create a forum for sharing technical expertise developed within university systems on behalf of their campuses and the national community."

The Cyberinfrastructure and policy components needed to develop and extend an end-to-end CI ecosystem include:

- Regional centers of excellence in CI based in one or a consortium of regional CI providers, and to create ongoing forums for building and coordinating regional CI support collaborations; and

- Funding mechanisms for CI that support the development of shared CI, shared expertise and shared training and education based on shared needs of many institutions in a region.

References
1. Monaco, G.E., McMullen, D.F., Huntoon, G., Leasure, J., Swanson, D., Neeman, H., Blake, J., Adams, K. The Role of Regional Organizations in Improving Access to the National Computational Infrastructure: A Report to the National Science Foundation, June, 2016.

Question 3 Other considerations (maximum ~1200 words, optional): Any other relevant aspects, such as organization, process, learning and workforce development, access, and sustainability, that need to be addressed; or any other issues that NSF should consider.

In a recent conference funded by the National Science Foundation (The Role of Regional Organizations in Improving Access to the National Computational Infrastructure) the following needs were identified:

- Overcoming barriers to accessing cyberinfrastructure at smaller colleges and universities and in disciplines at better resources campuses where adequate training and support may not be readily available.

- Promotion of cyberinfrastructure awareness and literacy.

- Bridging the gap between cyberinfrastructure providers and CI consumers.

- Incentives for sharing cyberinfrastructure resources across campuses.

- Sustainability - Organizations with local "customizations" are more sustainable in the long run, and are better tuned to evolve to meet future needs.

In our subsequent report to the National Science Foundation [1], condensed and published in a recent Educause bulletin [2], we noted that "Regional organizations such as regional optical networks and regional high-performance computing centers can serve as the connector between users and CI resources. Proximity to their campus members and familiarity with regional priorities and interests allow these organizations to focus on challenges and opportunities characteristic of the region, as well as to promote new capabilities and resources beyond individual campuses.

"One or more regional organizations can potentially serve as a CI hub and occupy a key strategic position between campuses and national

computational resources. With respect to computational- and data-intensive science and research, regional organizations have shared values and are poised to move the national science agenda forward through more effective education, outreach, coordination, and sharing of resources, including developing personal relationships to support and sustain initiatives.

"In order to do this most effectively, coordination across regional organizations is necessary to develop tactics to maximize adoption of successful activities, minimize duplication of effort, and gain a better understanding of what works. Some key areas where this work can occur include broadening access to the computational infrastructure, promoting CI awareness and literacy, developing a culture of bridging the gap, and sharing resources."

In particular, we highlighted the role of state and regional networks because they are “well positioned to help expand access to all components of the CI ecosystem, including computation, data storage, and expertise. They are self-sustaining and have connections to smaller colleges and universities.” A set of best practices were identified to assist regional networks to experiment with becoming cyberinfrastructure centers of excellence. The recommendations for these organizations enumerated in these reports [1,2] are to

- “Add personnel dedicated to computational education and outreach efforts. These are professionals who can communicate about both the science and the cyberinfrastructure. Their role is to enable researchers to more effectively and efficiently reach their research goals.

- “Actively seek out and communicate with researchers at smaller institutions who have unmet data-storage and/or computational needs. Regional network personnel are in a unique position to help these researchers—who often have limited time and limited local resources—navigate the CI terrain.

- “Identify and catalogue computational and data resources across their region. Even within campuses, computational and other network-reachable resources are often unknown to researchers and campus IT departments. An inventory of resources would be invaluable, not only for researchers but also for regional network organizations in order to provide more resources while keeping costs down.

- “Facilitate the sharing of computational and data resources across their region. Not all computational resources are appropriate for all research problems, and not all researchers have ready access to computational resources. Sharing of resources across campus boundaries will expand the availability of the most appropriate resources to researchers who need them.

- “Expand educational and meeting activities to include topics of interest to researchers with computational and data storage needs. Regional networks are geographic centers for CI and can convene researchers from larger and smaller campuses to learn from one another, to problem solve, and to obtain campus, state, regional, and national updates.

- “Coordinate efforts with one another (e.g., via The Quilt). State and regional networks have a long history of working together and assisting one another. The Quilt, the member organization of regional networks, currently has an initiative to strengthen researcher engagement and increase coordination efforts within the research and education networking community.

- “More closely coordinate with national activities like XSEDE Campus Champions, OSG, and ACI-REF. Regional networks can serve as a communication conduit regarding campus and regional infrastructure and requirements while sharing information about national computation resources with their campus communities.

- “Coordinate efforts with national research and education networks such as Internet2 and ESnet. These networks have similar goals to support academic research for the benefit of furthering scientific exploration.

- “Coordinate end-to-end performance troubleshooting. When problems occur, it is important to rapidly discover where things went awry and get them back on track. This can be done by actively implementing and promoting performance measurement tools such as perfSONAR. “

References

1. Monaco, G.E., McMullen, D.F., Huntoon, G., Leasure, J., Swanson, D., Neeman, H., Blake, J., Adams, K. The Role of Regional Organizations in Improving Access to the National Computational Infrastructure: A Report to the National Science Foundation, June, 2016.

2. Monaco, G.E., & McMullen, D.F. Improving Access to the National Computational Infrastructure: The Role of Regional Organizations. Educause Center for Analysis and Research (ECAR) Research Bulletin, 2016.

Consent Statement

- "I hereby agree to give the National Science Foundation (NSF) the right to use this information for the purposes stated above and to display it on a publically available website, consistent with the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode)."